What is claimed is:

1.   Method for recognizing text in a captured imagery, said method comprising the steps of:

     (a) detecting a text region in the captured imagery;

5   (b) adjusting said detected text region to produce a rectified image; and

     (c) applying optical character recognition (OCR) processing to said rectified image to recognize the text in the captured imagery.

2.   The method of claim 1, wherein said adjusting step (b) comprises the

10 step of (b1) computing a base line and a top line for a line of detected text within said detected text region.

3.   The method of claim 2, wherein said base line and said top line correlate substantially to horizontal parallel lines of a rectangular bounding box that is

15 fitted to said line of detected text.

4.   The method of claim 2, wherein said base line and said top line are estimated by rotating said line of detected text at various angles and then computing a plurality of horizontal projections over a plurality of vertical edge

20 projections.

5.   The method of claim 4, wherein said base line is selected that corresponds to a rotation angle that yields a steepest slope on a bottom side of one of said plurality of horizontal projections.

25

6.   The method of claim 4, wherein said top line is selected that corresponds to a rotation angle that yields a steepest slope on a top side of one of said plurality of horizontal projections.

30 7.   The method of claim 2, wherein said base line is selected comprising the steps of:

locating a plurality of bottom edge pixels, where each bottom edge pixel is located for each column in said rectangular bounding box;

rotating said plurality of bottom edge pixels through a series of angles around an initial estimated text angle for said line of detected text;

5   summing horizontally along each row; and

determining a baseline angle from a maximum sum of squared projections and determining a baseline position from a maximum projection.

8.   The method of claim 2, wherein said top line is selected comprising the

10   steps of:

locating a plurality of top edge pixels, where each top edge pixel is located for each column in said rectangular bounding box;

rotating said plurality of top edge pixels through a series of angles around an initial estimated text angle for said line of detected text;

15   summing horizontally along each row; and

determining a top line angle from a maximum sum of squared projections and determining a top line position from a maximum projection.

9.   The method of claim 2, wherein said adjusting step (b) further comprises

20   the step of (b2) computing a dominant vertical direction of character strokes for a line of detected text within said detected text region.

10.   The method of claim 9, wherein said dominant vertical direction computing step (b2) comprises the step of computing a plurality of vertical

25   projections over a plurality of vertical edge transitions after rotating said line of detected text in a plurality of degree increments.

11.   The method of claim 10, wherein said dominant vertical direction is selected that corresponds to an angle where a sum of squares of said vertical

30   projections is a maximum.

12.   The method of claim 1, further comprising the step of:

(b1) binarizing said detected text region prior to applying said OCR processing step (c).

13. The method of claim 12, further comprising the step of:

5　　(d) applying agglomeration processing subsequent to said OCR processing to produce the text in the captured imagery.

14. The method of claim 13, further comprising the step of:

(e) applying lexicon processing subsequent to said agglomeration

10　processing to produce the text in the captured imagery.

15. The method of claim 14, further comprising the step of:

(f) applying false text elimination processing subsequent to said lexicon processing to produce the text in the captured imagery.

15

16. Apparatus for recognizing text in a captured imagery, said apparatus comprising:

means for detecting a text region in the captured imagery;

means for adjusting said detected text region to produce a rectified

20　image; and

means for applying optical character recognition (OCR) processing to said rectified image to recognize the text in the captured imagery.

17. The apparatus of claim 16, wherein said adjusting means computes a

25　base line and a top line for a line of detected text within said detected text region.

18. The apparatus of claim 17, wherein said base line and said top line correlate substantially to horizontal parallel lines of a rectangular bounding

30　box that is fitted to said line of detected text.

19. The apparatus of claim 17, wherein said base line and said top line are estimated by rotating said line of detected text at various angles and then computing a plurality of horizontal projections over a plurality of vertical edge projections.

5

20. The apparatus of claim 19, wherein said base line is selected that corresponds to a rotation angle that yields a steepest slope on a bottom side of one of said plurality of horizontal projections.

10 21. The apparatus of claim 19, wherein said top line is selected that corresponds to a rotation angle that yields a steepest slope on a top side of one of said plurality of horizontal projections.

22. The apparatus of claim 17, wherein said adjusting means further
15 computes a dominant vertical direction of character strokes for a line of detected text within said detected text region.

23. The apparatus of claim 22, wherein said adjusting means computes said dominant vertical direction by computing a plurality of vertical projections over
20 a plurality of vertical edge transitions after rotating said line of detected text in a plurality of degree increments.

24. Method for recognizing text in a captured imagery having a plurality of frames, said method comprising the steps of:
25   (a) detecting a text region in a frame of the captured imagery;
    (b) applying optical character recognition processing (OCR) to said detected text region to identify potential text for said frame; and
    (c) agglomerating the OCR identified potential text over a plurality of frames in the captured imagery to recognize the text in the detected text
30 region.

25. The method of claim 24, wherein said agglomerating step (c) comprises the step of updating an agglomeration structure with said OCR identified potential text of a current frame.

5 26. The method of claim 25, wherein said updating step comprises the step of (c1) finding correspondence between a text region of said agglomeration structure with a text region of said current frame.

27. The method of claim 26, wherein said updating step further comprises
10 the step of (c2) finding character-to-character correspondence for each pair of overlapping lines between said text region of said agglomeration structure with said text region of said current frame to find one or more character group pairs.

28. The method of claim 27, wherein said updating step further comprises
15 the step of (c3) updating said one or more character group pairs.

29. The method of claim 28, wherein said updating step further comprises the step of (c4) marking text in said agglomeration structure that is not in said current frame as a deletion.

20 30. The method of claim 29, wherein said updating step further comprises the step of (c5) marking text in said current frame that is not in said agglomeration structure as an insertion.

25 31. The method of claim 25, further comprising the step of:
(d) outputting said text in the detected text region after each processed frame.

32. The method of claim 25, further comprising the step of:
30 (d) outputting said text in the detected text region only when a change is detected as to said text of said captured imagery.

33.   The method of claim 25, further comprising the step of:
        (d) outputting only said text within said agglomeration structure when said text is not detected in a current frame.

34.   Apparatus for recognizing text in a captured imagery having a plurality of frames, said apparatus comprising:
        means for detecting a text region in a frame of the captured imagery;
        means for applying optical character recognition processing (OCR) to said detected text region to identify potential text for said frame; and
        means for agglomerating the OCR identified potential text over a plurality of frames in the captured imagery to extract the text in the detected text region.

35.   The apparatus of claim 34, wherein said agglomerating means updates an agglomeration structure with said OCR identified potential text of a current frame.

36.   The apparatus of claim 35, wherein said agglomerating means finds correspondence between a text region of said agglomeration structure with a text region of said current frame.

37.   The apparatus of claim 36, wherein said agglomerating means further finds character-to-character correspondence for each pair of overlapping lines between said text region of said agglomeration structure with said text region of said current frame to find one or more character group pairs.

38.   The apparatus of claim 37, wherein said agglomerating means further updates said one or more character group pairs.

39.   The apparatus of claim 38, wherein said agglomerating means further marks text in said agglomeration structure that is not in said current frame as a deletion.

40.   The apparatus of claim 39, wherein said agglomerating means further marks text in said current frame that is not in said agglomeration structure as an insertion.

41.   The apparatus of claim 35, further comprising:

        means for outputting said text in the detected text region after each processed frame.

42.   The apparatus of claim 35, further comprising:

        means for outputting said text in the detected text region only when a change is detected as to said text of said captured imagery.

43.   The apparatus of claim 35, further comprising:

        means for outputting only said text within said agglomeration structure when said text is not detected in a current frame.